Comparison of echogram measurements against data expectations and assumptions for distinguishing seafloor substrates

Mark Zimmermann (contact author)
Christopher N. Rooper
Email address for M. Zimmermann: Mark.Zimmermann@noaa.gov
National Marine Fisheries Service
Alaska Fisheries Science Center
7600 Sand Point Way NE, Bldg. 4
Seattle, Washington 98115-6349

Abstract—Defining types of seafloor substrate and relating them to the distribution of fish and invertebrates is an important but difficult goal. An examination of the processing steps of a commercial acoustics analyzing software program, as well as the data values produced by the proprietary first echo measurements, revealed potential benefits and drawbacks for distinguishing acoustically distinct seafloor substrates. The positive aspects were convenient processing steps such as gain adjustment, accurate bottom picking, ease of bad data exclusion, and the ability to average across successive pings in order to increase the signal-to-noise ratio. A noteworthy drawback with the processing was the potential for accidental inclusion of a second echo as if it were part of the first echo. Detailed examination of the echogram measurements quantified the amount of collinearity, revealed the lack of standardization (subtraction of mean, division by standard deviation) before principal components analysis (PCA), and showed correlations of individual echogram measurements with depth and seafloor slope. Despite the facility of the software, these previously unknown processing pitfalls and echogram measurement characteristics may have created data artifacts that generated user-derived substrate classifications, rather than actual seafloor substrate types.

Manuscript submitted 4 February 2008.
Manuscript accepted 28 March 2008.
Fish. Bull. 106:293–304 2008).

Marine natural resource managers must define essential fish habitat (EFH) for federally managed, commercially exploited species (Federal Register, 2002) but the best method for fulfilling this mandate across the vast area and significant depths of the U.S. Exclusive Economic Zone remains unknown. A successful acoustic method for determining EFH would be of great benefit, because single-beam seafloor echosounder reflections are collected simultaneously with fish density estimates during National Marine Fisheries Service (NMFS) stock assessment bottom trawl surveys in the Gulf of Alaska (~800 stations among 320,000 km^2, ≤1000 m depth) and the Aleutian Islands (~400 stations among 67,000 km^2, ≤500 m depth). Therefore we conducted an acoustic analysis on data from a small portion from one survey in order to determine if there was a direct correlation between substrate classes or echogram measurements with species abundance.

We tested a widely used, proprietary software package (vers. 3.30, QTC IMPACT™), developed by the Quester Tangent Corporation (QTC, Sidney, British Columbia, Canada), to resolve the echosounder reflections into substrate types for comparison with the survey trawl catch data to determine whether there was a correlation or relationship between seafloor substrate classes and fish-density. This software produces 166 proprietary unitless echogram measurements (EMs) on the first seafloor echo for an internal principal components analysis (PCA), and then uses the first three principal components (PCs), generally accounting for more than 95% of the covariance (Ellingsen et al., 2002; Legendre et al., 2002) in *K*-means clustering, for dividing the first seafloor echoes into acoustically distinct substrate types.

Our initial efforts with *K*-means clustering indicated that a solution of any particular number of classes was not much better than other solutions (e.g., four versus five substrate classes), and therefore the 166 EMs were analyzed to determine if they could be used in another analysis for resolving substrate types. Although the general manner in which the 166 EMs, or the data, are acquired, processed, and divided into substrate classes by QTC software has been well reported in the literature, many specific details are lacking and it was therefore not clear what these 166 EMs represent.

To investigate an acoustic method for determining EFH we described the specific details of the processing method that QTC software follows, focusing on potential pitfalls and advantages for the user. We report on new findings based on some simple data explorations on the 166 EMs from echosounder data collected during a 2003 NMFS research cruise; and our findings are corroborated with four data sets collected independently from other agencies on other ships. In this analysis we checked the assumption that these 166 EMs have the same scale or range as that normally used in PCA, and the assumption that these 166 EMs are derived from the first echo only. Because both of these assumptions are typically presumed to be correct for this type of acoustic analysis, these findings may be of use for interpreting seafloor substrate classifications for determining EFH.

Materials and methods

Data collection and conversion

Data were collected in the *.raw format from a 38-kHz Simrad single-beam echosounder on the FV *Gladiator* during the 2003 NMFS bottom trawl survey in the Gulf of Alaska (Table 1). The transducer gain was 24.5 dB, transmit power was 1500 W, beam angle was 9°, pulse length was 4.096 ms, and the sampling interval was 1.024 ms. These Simrad files were calibrated in EchoView® (vers. 3.30.60.05, SonarData Pty. Ltd, Hobart, Tasmania, Australia), and short (~1.5 km, ~1440 pings) seafloor sections corresponding to 15-minute duration bottom tows conducted at 1.54 m/s (3 knots) were exported into binary files by using the EchoImpact export module for import into QTC IMPACT™. This EchoImpact export module was specifically designed by the two companies to convey acoustic data in an appropriate format from EchoView to QTC IMPACT. We also examined EMs recorded directly by QTC VIEW™ (QTC, Sidney, British Columbia, Canada), without any prior EchoView® processing, at preset gains (ping intensities or amplitudes) by four external research cruises: the Alaska Department of Fish and Game (ADFG RV *Resolution* 2003 cruise), the Canadian Department of Fisheries and Oceans (Canadian coast guard ship RV *John P. Tully* 2002 cruise, and the RV *Pallasi* 2004 cruise), and the New Zealand National Institute of Water and Atmosphere (RV *Rangithi* 1999 cruise) (Table 1).

Gain settings

In automated seafloor echo-processing systems, there may be a mismatch between the seafloor echo strength (gain) and the ability of the processing system to identify the abrupt rise or spike that represents the beginning of the seafloor reflection. It is necessary to adjust the gain setting such that the inflection point can be distinguished from the earlier portion of the echo, which is the water column above the seafloor. Therefore several postprocessing gain settings in QTC IMPACT were applied to subsets of the NMFS 2003 FV *Gladiator* data sets in order to maximize the number of pings strong enough for automatic bottom detection (or bottom picking) and to minimize the number of pings that would be too strong for the dynamic range of 96 dB of sound that QTC software can process. Otherwise, louder portions of pings would have had to have been automatically decreased to 0 dB, a process known as clipping, and quieter portions of pings would have had to be automati-

Table 1

Details of echosounder research cruises and of the data collected from these cruises from 1999–2004 for a study to determine whether there is a correlation between echogram measurements and species abundance. Codes for agencies: NMFS, National Marine Fisheries Service; ADFG, Alaska Department of Fish and Game; NIWA, New Zealand National Institute of Water and Atmosphere; DFO, Canadian Department of Fisheries and Oceans. QTC VIEW and QTC IMPACT are software products from QTC (Quester Tangent Corporation), Sidney, British Columbia, Canada. Reference depth is the depth for which depth-related changes in echo signal protraction were corrected. Stacks are groups of five echoes or acoustic returns that were summed together for analysis. Fully collinear echogram measurements were the number of values out of a possible 166 that caused the variance-covariance matrix determinant in a principal components analysis to be zero.

Vessel	Year	Location	Agency	Echo sounder	kHz	Sampling (Hz)	QTC VIEW vers.	QTC IMPACT vers.	Reference depth (m)	Depth range (m)	Stacks	Number of fully collinear echogram measurements
FV *Gladiator*	2003	Gulf of Alaska	NMFS	Simrad	38	977	none	3.30	50	25–100	14,432	155
FV *Gladiator*	2003	Gulf of Alaska	NMFS	Simrad	38	977	none	3.30	150	100–200	3598	155
RV *Resolution*	2003	Gulf of Alaska	ADFG	Biosonics 101	120	20,000	3.25	3.3	100	66–136	3680	152
RV *Rangithi*	1999	New Zealand	NIWA	Simrad EA501P	200	20,000	4	2	10	4–15	736	148
RV *Pallasi*	2004	B.C., Canada	DFO	Furuno	50	20,000	4	3.4	12	3–34	736	155
RV *John P. Tully*	2002	B.C., Canada	DFO	Simrad	50	20,828	5	3.20	50	64–132	727	152

cally increased to –96 dB. Another convenience of the QTC software was that abnormally weak echoes could be eliminated by specifying a minimal signal strength, and this was set to be equal to 25% of the maximum permissible amount (0 dB).

Bottom picks

After importing the recorded echoes into the software at an appropriate gain setting, another method for further improving the signal-to-noise ratio would be to assemble a stack of successive echoes, presumably from the same substrate, and average the echo stack into a single echo (Pace and Ceen, 1982). For QTC IMPACT, a minimum stack size of five pings was recommended, which corresponded to 2.5 seconds or 3.85 m traveled at 1.54 m/s, for a theoretical yield of 288 stacks per trawl path. The strong, positive benefits of stacking were dependent on correctly aligning events with the successive echoes, and therefore on the software's interpretation of the seafloor inflection point. Although this data check is not mentioned in the literature, it is a critical part of the process, because all measurements start at the seafloor inflection point. We examined every bottom pick for appropriate placement, as recommended by QTC guidelines. This process determined that the bottom pick was not interpolated between sample intervals, and that the natural variability of the depth among a group of pings would be the distance sound travels during half the sampling interval (0.768 m).

Generating echogram measurements

Once the bottom pick had been located, an automatic determination of the length or extent of any seafloor echo was difficult because rough, steep, soft, and deep areas have longer reflections than smooth, flat, hard, and shallow areas. The QTC IMPACT software uses 256 sound samples of vertical time intervals, or recorded sound intensity within a ping, surrounding the bottom pick. Starting at the bottom pick, five samples (representing the water column) were taken above the seafloor inflection point and 251 samples were taken below the start of the first seafloor reflection (representing the seafloor). If the echograms contained fewer than 251 time intervals below the start of the first seafloor reflection, the last sample was repeated (padded) as many times as needed until the 251 sample requirement was fulfilled. The QTC software then generated 166 EMs for each stack with reference to a specific depth such that depth-related changes in signal protraction were corrected.

Optimum substrate classification

Organizing the echogram measurements along a continuum of measurements or grouping them into a number of acoustically distinct substrate classes is the final step in the process. Ideally this step would identify substrate qualities of importance to EFH species, such that researchers could infer essential fish habitat from substrate types and use this information for better resource management. The QTC method first uses continua by performing PCA on the 166 EMs and retaining the first three PCs for plotting the location of each stack in three-dimensional space. Then it is up to the user to determine the optimum number of substrate classes on the basis of the *K*-means clustering of the first three PCs.

Examination of the data

Because the algorithms for producing the 166 EMs are proprietary, the data values produced by the 166 EMs were exported and viewed in a text editor, which showed that the data values for each stack were displayed as seven decimal-place numbers in four columns, underneath a stack header. In order to resolve the possible complications of having a single set of 166 EMs for five different stacked pings, a single ping was exported from EchoView® and imported into QTC IMPACT five times, to create one stack of identical pings. The four columns of 166 EMs were reformatted into a single column in a spreadsheet and an examination of the data revealed that the EMs from this single, repeated ping were occurring in five groups (von Szalay, 1998).

Variability and covariance of echogram measurements

Simple data checks, such as checks of averages, variances, minima, and maxima, enabled us to describe each data set and determine the range or scale of each EM. The variance between EMs, or the covariance, was derived to determine the amount of collinearity among the EMs.

Correlation of echogram measurements with depth

The correlation between each EM versus depth was determined from each of the data sets. This simple analysis, which could provide some useful diagnostics, has not been reported in any of the literature.

Angle of incidence

The angle of echosounder seafloor reflections has a potentially confounding influence on any depth-correlation analysis, because the rate of change of depth and slope vary together. In general, QTC and similar products should be used to analyze normal (90°) incident reflections (see Pace and Ceen, 1982; Orlowski, 1984), and it is expected that severe departures from normality would cause analytical failures. The influence of non-normal (<90°) reflections could not be formally examined in our study because of a lack of knowledge about cross-track slope, vessel pitch and roll, and interactions between seafloor angle and vessel angle. However, more single-beam data were analyzed from the FV *Gladiator* in 2005 at small study sites in the Aleutian Islands that had been groundtruthed with video and multibeam sonar equipment in 2004, such that the substrate types and seafloor slopes were known (Rooper and Zimmermann, 2007).

Assumption 1: scale of measurements

Although as a group the 166 EMs ranged between zero and one (Legendre et al., 2002), this was tested for each of the 166 EMs; and EMs fully extending across this range would indicate that the data had been standardized for proper PCA (Manly, 1994; Legendre and Legendre, 1998). PCA was also conducted (S-Plus, vers. 6.1, Insightful Corp., Seattle, WA) independently to ensure that the QTC PCA results could be reproduced.

Assumption two: first echo

Although it is widely reported or implied in the literature that QTC IMPACT software analyzes only the first echo, that conclusion is not strictly correct. QTC IMPACT analyzes the first 251 sound samples beneath the bottom pick, and it is up to the user to ensure that this is a meaningful window. The relationship between the 251 samples and the analysis depth range is directly related to half of the sample interval preset in the echosounder;

$$\textit{Analysis depth range}\ (\mathrm{m}) = \textit{Sample interval}\ (\mathrm{s/sample}) \times (0.5) \times (1500\ \mathrm{m/s}) \times 251\ \text{samples},$$

where 1500 m/s = the approximate speed of sound through seawater.

This relationship was tested to determine how well the 251 sample size corresponded with the first echo in the NMFS 2003 FV *Gladiator* data sets and with the external data sets collected by other agencies from other vessels.

Results

Gain settings

Determining the proper gain setting for the NMFS 2003 FV *Gladiator* data sets was a time-intensive process because a wide range of gains needed to be applied and weak or bad data had not yet been identified. We determined that a gain setting of –18 dB would be appropriate for shallow sites (25–100 m, 68 sites, 97,119 pings) and a gain setting of –17 dB would be appropriate for deep trawl sites (100–200 m, 19 sites, 25,110 pings). Several additional sites with too many weak pings (>50%), an indication of bad data, were identified and eliminated from processing at this stage. Although changing the gain setting within QTC IMPACT by 0.5 dB was equivalent to changing the gain setting by 1.0 dB in Echo-View®, use of the gain adjustment within QTC IMPACT was far more convenient for adjusting the echo signal strength to be within the required 96-dB range, and for identifying bad data.

Bottom picks

In the NMFS 2003 FV *Gladiator* data sets, there were 72,296 shallow (25–100 m) pings with bottom picks including 3961 that were clipped, and there were 18,021 deep (100–200 m) pings with bottom picks including 1106 of those that were clipped. Thus there was a greater than 70% success rate in bottom picking and approximately 6% clipping among both data sets, indicating that the gain settings were appropriate. Each bottom pick was inspected and found to occur anywhere between the base and the tip of the peak—in the general region of the seafloor location. Thus QTC IMPACT software did an excellent job of bottom picking in the NMFS 2003 FV *Gladiator* shallow and deep data sets.

Generating echogram measurements

The NMFS 2003 FV *Gladiator* shallow data set yielded 14,432 stacks (of five pings) and the deep data set yielded 3598 stacks (of five pings); odd lots of fewer than five pings were not included in stacks, and therefore the total number of stacks was slightly less than one fifth of the total number of pings with bottom picks. Padding was required at all shallow sites for all of the stacks, and padding was required at 16 of 19 deep sites on a total of 3112 stacks. A reference depth of 50 m was used for the shallow sites and 150 m was used for the deep sites. The EMs for the shallow sites were combined into a single data set for PCA and *K*-means clustering. The process was repeated for the EMs from the deep sites.

Optimum substrate classification

The *K*-means clustering of the first three PCs indicated that a solution of any specific number of acoustically derived substrate classes would not explain much more of the variance than other solutions. Therefore the data processing was repeated several times to check for errors that may have influenced the results. The main focus was on gaining a better understanding of the numbers that were being created and processed with PCA and *K*-means clustering, and on exploring factors that may have affected the EMs.

Examination of the data

Examinations of the spreadsheets of EMs from the NMFS 2003 FV *Gladiator* shallow and deep data sets, and the four externally collected data sets, showed the same groups as those revealed by the examination of the stack of the single pings repeated five times; EMs 1–23, EMs 24–39, EMs 40–70, EMs 71–101, and EMs 102–166. Across all data sets, the EMs in the first (EMs 1–23) and fifth (EMs 102–166) groups were, in general, highly correlated with their neighbors (e.g., EM 22 versus 23, Fig. 1). In the second group of EMs (EMs 24–39), EM 31 was the sum of EM 32 through EM 39, each of which were fractions of 256 (e.g., 1/256, 2/256). Among the 31

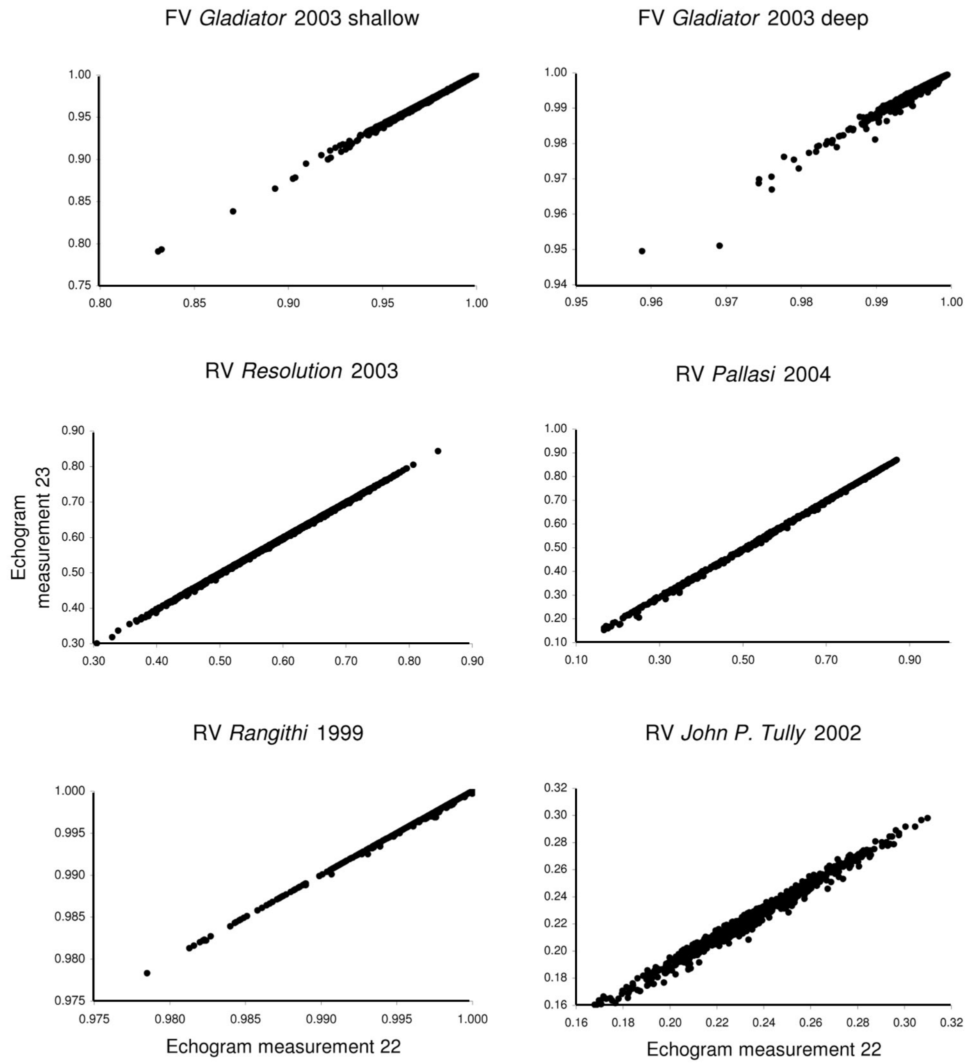

Figure 1

Significant linear regressions ($P<0.0005$) between the unitless Quester Tangent Corporation echogram measurements 22 and 23 for the shallow (n=14,432) and deep (n=3598) FV *Gladiator* 2003 data sets, the RV *Resolution* 2003 cruise (n=3680), the C.C.G.S. RV *John P. Tully* 2002 cruise (n=727), the RV *Pallasi* 2004 cruise (n=736), and the RV *Rangithi* 1999 cruise (n=736). Note the different scales on the x axes.

EMs in each of the third (EMs 40–70) and fourth (EMs 71–101) groups, there were 16 original EMs grouped into 15 succeeding sums, with the final EM (EM 40 in the third group and EM 71 in the fourth group) being the sum of the whole group (Fig. 2). In the third group, EM 40, which was always 1.0, was created by the sum of EM 41 and EM 56, which are complements to each other and therefore are entirely dependent.

Variability and covariance of echogram measurements

There were several unusual observations of variance or covariance among the 166 QTC IMPACT-generated EMs. Three EMs—16, 31, and 40—never varied (were always 1.0000000 or 0.9999999) and it is presumed these are the same three that Legendre et al. (2002) stated never varied from one. Additionally, three EMs from the RV *Rangithi* 1999 cruise data were always zero. Although it is difficult to describe all the dependent relationships (the sums and correlations) between the remaining 160 to 163 EMs, it is much simpler to note that among the six data sets, a range of 148 to 155 EMs (Table 1) were fully collinear (causing the variance-covariance matrix determinant to be zero; Neter et al., 1990). Among the eight to 12 EMs within each data set that were not fully collinear, four to

Initial sums	Secondary sums	Tertiary sums	Final sum
70 + 69 = 68			
67 + 66 = 65			
	68 + 65 = 64		
63 + 62 = 61			
60 + 59 = 58			
	61 + 58 = 57		
		64 + 57 = 56	
55 + 54 = 53			
52 + 51 = 50			
	53 + 50 = 49		
48 + 47 = 46			
45 + 44 = 43			
	46 + 43 = 42		
		49 + 42 = 41	
			56 + 41 = 40 = 1.000000

Figure 2

Schematic diagram depicting a sequence (from left to right) of 15 sums from 16 original unitless echogram measurements—a sequence of sums that leads to a final sum of one. The left column shows how 16 original echogram measurements are summed into eight new echogram measurements. The next column shows how these eight sums are summed into four new echogram measurements. The third column shows how these four sums are summed into two new echogram measurements. The fourth column shows how these two sums are summed to produce a final echogram measurement of one.

seven EMs had variance inflation factors >10 (a general threshold indicating high correlations but not full collinearity with the remaining variables; Neter et al., 1990), leaving only three to six relatively independent EMs in each data set.

Correlation of echogram measurements with depth

There was a significant relationship between depth and some EMs in all data sets (Fig. 3). This relationship translated into significant relationships (LOESS curve fits) between PC1 and PC2 versus depth for all six data sets (F-tests, $P<0.001$), indicating that depth has a direct influence on the QTC substrate classification.

Angle of incidence

At the Aleutian Islands groundtruth site (FV *Gladiator* in 2005; Rooper and Zimmermann, 2007), there were significant linear correlations ($P<0.05$) between slope and EMs for the most common substrate classes of sand-boulder ($n=368$ video observations), sand-sand ($n=351$), and bedrock-boulder ($n=259$), even when the analyses were restricted to low (<5°) slopes (von Szalay, 1998; von Szalay and McConnaughey, 2002). The influence of slope resulted in EMs that were equivalent among different substrates and different slopes. For example, EM 1 on a substrate of bedrock-boulder at 1° slope was equivalent to EM 1 on sand-sand substrate at 4.1° slope, and equivalent to EM 1 on sand-boulder substrate at 4.9° slope (Fig. 4), illustrating how easily substrates can be misclassified at low slopes.

Assumption one: scale of measurements

The S-Plus version of PCA, conducted after eliminating invariant EMs (16, 31, and 40), confirmed that the QTC IMPACT method of PCA does not use any additional data ranging or standardization. PCA performed in S-Plus with standardization (subtraction of mean,

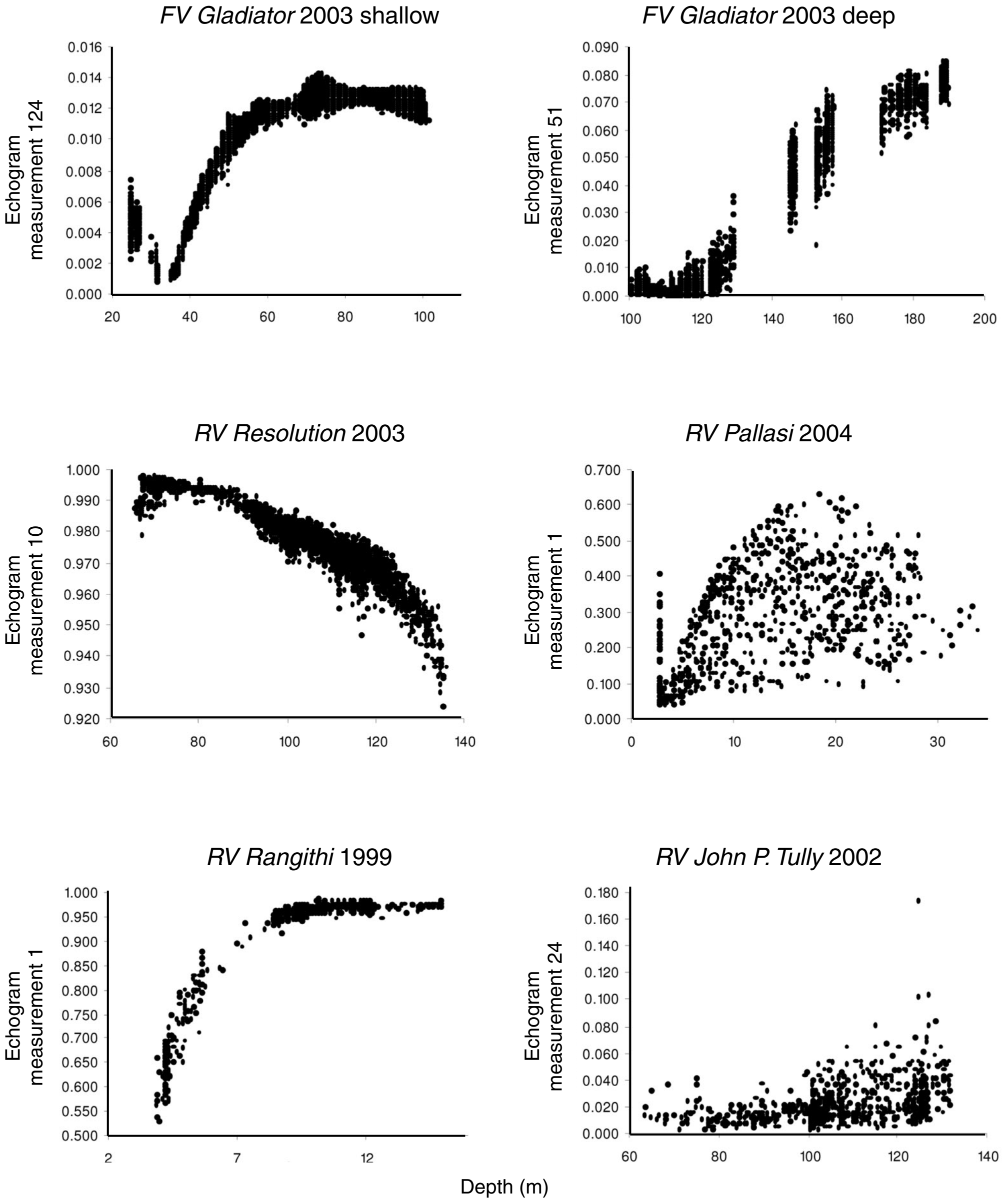

Figure 3

Significant relationships between QTC (Quester Tangent Corporation, Sidney, British Columbia, Canada) IMPACT™ generated echogram measurements and depth for the shallow and deep FV *Gladiator* 2003 data sets, the RV *Resolution* 2003 cruise, the C.C.G.S. RV *John P. Tully* 2002 cruise, the RV *Pallasi* 2004 cruise, and the RV *Rangithi* 1999 cruise. Note the different scales on the x axes.

division by standard deviation; Manly, 1994; Legendre and Legendre, 1998) revealed that more PCs were required to explain the same total amount of variance as the QTC IMPACT PCA method (Fig. 5). Thus the lack of standardization within the QTC IMPACT PCA method can have a strong effect. For example, across all six data sets, some EMs such as EM 15 were always large ($\geq$0.938), some such as EM 166 were always small ($\leq$0.006), and some such as EMs 1 and 102 generally had larger ranges and were more variable. Within the first group of EMs (1–23), EM 1 was always $\leq$ EM 2, EM 2 was always $\leq$ EM 3, etc., up to EM 16, which was always $\geq$ EM 17, which was always $\geq$ EM 18, etc., up to EM 23. Thus these variables have constricted ranges which can affect PCA. Additionally, the strong correlations between neighboring variables within the first (EMs 1–23) and fifth (EMs 102–166) groups indicate that these EMs are either measuring nearly the same echo component, or that EMs are based on neighboring EMs. Without standardization of these correlated EMs, the amount of variance explained by the first three PCs is artificially inflated. The inclusion of sums of variables in a PCA, such as the 15 sums of variables in the third (EMs 40–70) and fourth (EMs 71–101) groups, also artificially inflates the amount of variance explained in a PCA. Inclusion of a variable that is a complement of another variable (EM 41 or EM 56) in a PCA does not improve or impair the results and one of these complements could be excluded, along with the three invariant variables.

Assumption two: first echo

The 251 sampling envelope used with the QTC IMPACT software may be a mismatch for the actual length of the first echo. In the FV *Gladiator* 2003 shallow and deep data sets (both <200 m), the 251 sampling intervals of 977 Hz or 0.001024 s translated into an excessive and unnecessary 192.8 m analysis depth range below the start of the first seafloor reflection, a fact not realized during the collection of the echosounder data. However, the recording of most of the *.raw files were truncated before the full 192.8 m distance, before any second echo, and also before the full 251 samples; therefore most of the echoes needed padding (extended repetition of the last sound sample in the echo). The second echo was recorded in the *.raw file and it fell within the 251 sample requirement in 19 of the 68 shallow sites (25–100 m) and four of the 19 deep sites (100–200 m). These same pings were exported from EchoView® to QTC IMPACT with and without the second echo, and our analyses demonstrated that QTC IMPACT treated the second echo as if it were part of the first echo. The accidental inclusion of the seafloor spike from this second echo reduced the values of the first 23 EMs (except EM 16) and had the greatest effect when the seafloor spike of the second echo occurred at the edge of the export window, where it was repeated to fulfill the 251 sample requirement (see Haul 206, Fig. 6). There was less of an effect when more of the second seafloor spike was included, so that the padded value was of lower sound intensity (see Haul 124, Fig. 6). Thus significant differences in some of the echogram measurements can be created for the exact same substrate type if users are not careful about ensuring that the 251 sample window of QTC IMPACT matches up well with the first echo length.

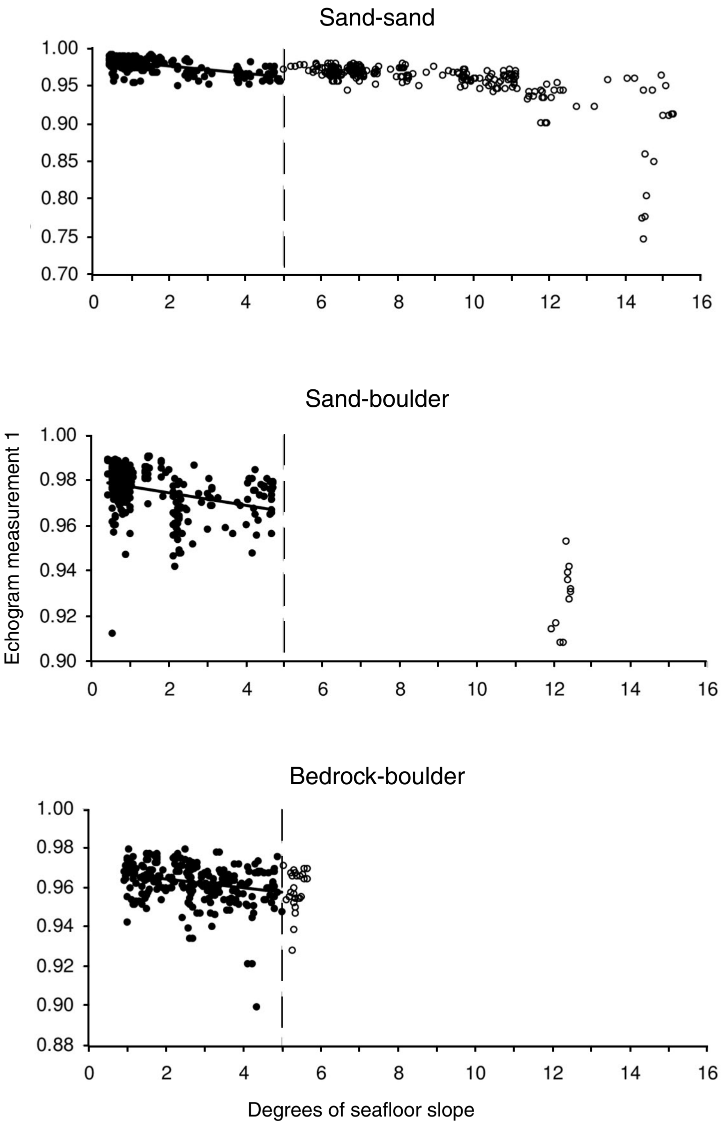

Figure 4

Significant linear regressions ($P<0.05$) between low seafloor slopes <5° (●) and QTC (Quester Tangent Corporation, Sidney, British Columbia, Canada) IMPACT™ echogram measurement 1, with observations from greater slopes (○) excluded from the regression analysis, for the three most common substrate types found during the FV *Gladiator* 2005 cruise. Only data values to the left of the 5° mark (dashed vertical line) were included in the regression analysis.

Discussion

Although this analysis demonstrated that there are several strong advantages (gain adjustment, bottom picking, bad data exclusion, and stacking) in using the partially automated echogram classifying software (QTC IMPACT), there are also several potential pitfalls (dependencies among the 166 EMs, lack of standardization, correlation with depth, influence of seafloor slope, and mismatch between 251 sample intervals versus first echo length), such that it does not function as users would expect for distinguishing substrate types. There are also several processing steps within QTC IMPACT, such as repeating the last sample of short pings (padding), reducing the strength of sections of pings that are too strong, and increasing the strength of sections of pings that are too weak, all of which may affect analyses. The proprietary nature of the software and the internal processing steps have discouraged user criticism and examination of the QTC IMPACT generated data (Kloser et al., 2001). QTC IMPACT users should export, format, and carefully examine their 166 EMs before substrate classification in order to catch user-generated mistakes, such as accidentally including a second echo, and to identify and remove any constant or collinear EMs. After analysis of the EMs, users may be able to reduce the 166 EMs to fewer than 10 without any loss of information, and compare these against depth, slope, and substrate types, if known, for further data-checking. The first assumption—that the scale or range of the data were appropriate for PCA—was disproved, and users may want to consider whether standardizing is appropriate for their data. The second assumption—that QTC IMPACT only uses the first echo—is not necessarily true and published QTC IMPACT substrate classes may have been differentiated by the presence or absence of all or part of the second echo.

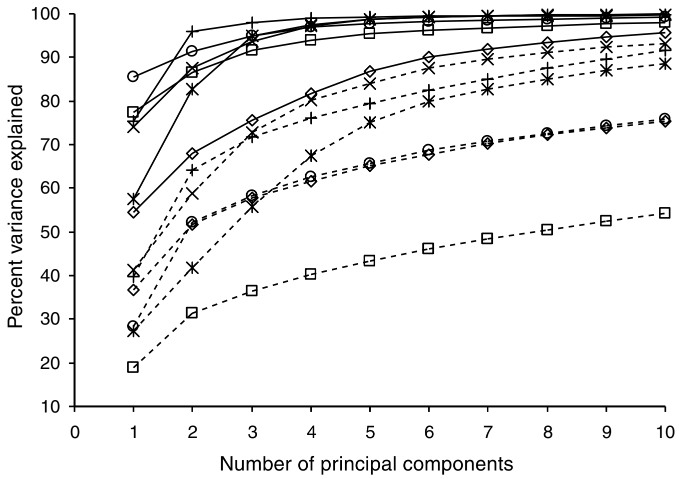

Figure 5

Percent of variance explained in principal components analysis (PCA) conducted on the echogram measurements produced by QTC (Quester Tangent Corporation, Sidney, British Columbia, Canada) IMPACT$^{\text{TM}}$ software when the data were not standardized (QTC method ———) and when the data were standardized (textbook method - - - -) for the RV *Resolution* 2003 cruise (□), the C.C.G.S. RV *John P. Tully* 2002 cruise (◇), the RV *Pallasi* 2004 cruise (○), the RV *Rangithi* 1999 cruise (+), and the NMFS FV *Gladiator* 2003 shallow (×) and deep (✳) cruises.

Optimum substrate classification

The inability to determine an optimum number of substrate classes for the shallow and deep FV *Gladiator* 2003 data sets is a common problem in seafloor substrate analysis and is not a critique of the particular *K*-means method within QTC IMPACT software. Our echosounder data could have been too noisy, too coarse, too affected by sea-state or seafloor slope, or our trawl sites could have been too variable or too constant for determining substrate classes. Instead, our results, with corroborations from independent data sets, indicated the importance of analyzing the echogram measurements before any PCA and *K*-means analysis so that depth-related and slope-related errors, second echo or echo envelope errors, and variable range or collinearity errors could be caught. The pros and cons of the QTC IMPACT method of *K*-means partitioning have already been thoroughly discussed. It was criticized by Legendre et al. (2002) who offered a new *K*-means method based on Euclidean distance. Preston and Kirlin (2003) responded by defending and elaborating on their *K*-means clustering method, which is based on Mahalanobis distance, and citing successful QTC IMPACT substrate-typing projects (Anderson, 2001; Morrison et al., 2001; Anderson et al., 2002; Ellingsen et al., 2002). Legendre (2003) offered additional criticism of the QTC IMPACT *K*-means clustering method and added that the QTC IMPACT method of clustering, based on only the first three principal components (PCs), was strongly influenced by depth, since his PC1 was strongly related to depth, as opposed to a solution that would include a greater number of PCs. Clearly the *K*-means method for distinguishing substrate classes is important, and strongly linked to the data values that feed into it.

Examination of the data and the variability and covariance of echogram measurements

This exploration of the 166 EMs provides the first description of the QTC IMPACT data set that is used for seafloor substrate classification. Before this description, only three EMs were known to be invariant and the rest were highly collinear (Legendre et al., 2002). These EMs were known to carried limited information and were highly redundant (Ellingsen et al., 2002). Researchers collecting data directly into QTC VIEW, such as corroborating data from different agencies, did not have the additional processing step of importing the data from EchoView®, and were probably unaware of the 96 dB dynamic range required for QTC IMPACT software. Therefore the effect of clipping portions of pings that were too loud, increasing the sound level of ping portions that were too quiet, or adding or subtracting a constant amount of sound to entire ping data sets (gain adjustment), was not widely reported in the literature. Only Anderson et al. (2002) mentioned experimenting with

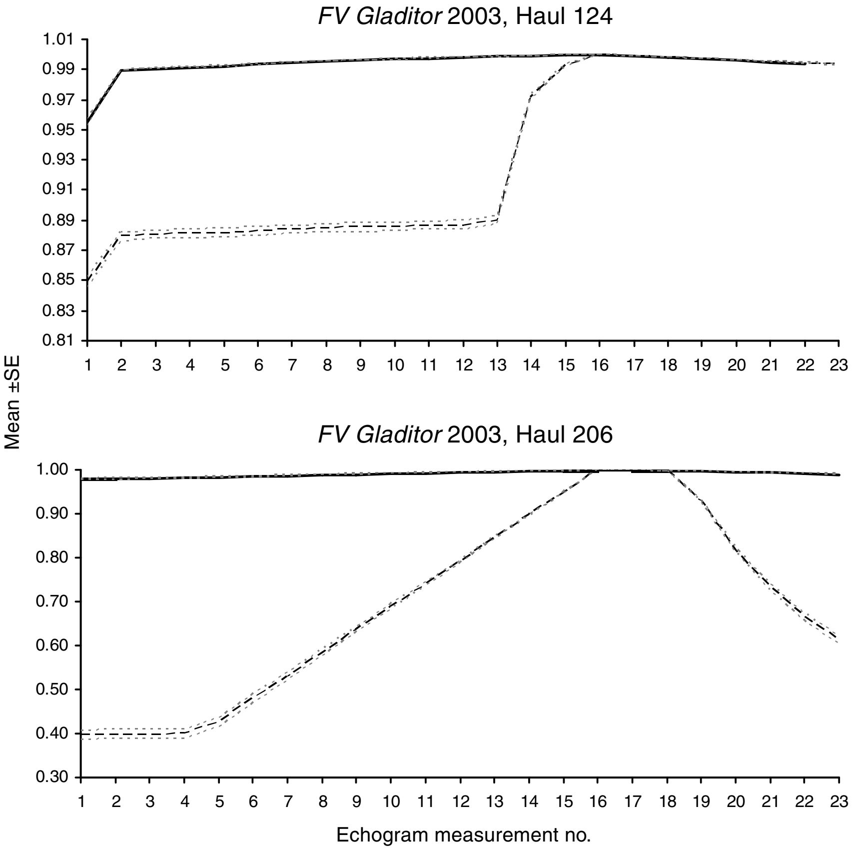

Figure 6

Relationship between the mean of the first 23 echogram measurements (± standard error [SE]) for two hauls when part of the second echo is included (- - - -) or excluded (———) in the QTC (Quester Tangent Corporation, Sidney, British Columbia, Canada) IMPACT^{TM} analysis.

different gain settings and different reference depths. Users have also been unaware of the 251 sample requirement, or the effects of repeating the last sound sample (padding) in order to fulfill this requirement. Ellingsen et al. (2002) mentioned that truncated acoustic reflections resulted in a loss of some of the 166 EMs. Perhaps results from their field work resulted in modification to the QTC IMPACT software so that the last acoustic sample was repeated (a process known as padding).

Correlation of echogram measurements with depth

The influence of depth on the EMs, and on the resulting PCs, may be due to improper echosounder calibration or improper depth-correction in QTC IMPACT, rather than to true variation in substrate types. It is not possible for users to determine the origin of the depth influence. Although the reference depth is supposed to compensate for the signal-protraction of pings of different depths within a data set, none of the QTC IMPACT studies in the literature have actually checked to determine if such compensation occurs. It has been reported in the literature that the QTC IMPACT-generated PC1 is correlated with depth (Legendre, 2003) and that QTC IMPACT-generated substrate classes are sometimes correlated with depth (Anderson et al., 2002). As with our findings, depth biases were also reported for the E1 (roughness) and E2 (hardness) measurements made by RoxAnn^{TM} (Sonavision, Aberdeen, Scotland, U.K.) bottom-typing software, as determined by a more thorough study with careful seafloor groundtruthing (Kloser et al., 2001).

Angle of incidence

Any potential effect due to impact angle of echogram reflection, which is a combination of seafloor slope and vessel motion, is not widely addressed in the literature. Anderson (2001) used QTC VIEW^{TM} to distinguish among substrates on steep slopes, some of which appear to be as steep as 45° (see Anderson, 2001, Figs. 4 and 5), whereas von Szalay (1998) and von Szalay and McConnaughey (2002) reported that seafloor slopes exceeding 5° to 8° caused significant substrate misclassification. In Ellingsen et al. (2002), effects from vessel motion may have been reduced or eliminated by working in calm seas. Our analysis of individual EMs revealed the mechanism whereby substrates may be misclassified in areas with slope and we would suggest that there is greater sensitivity with QTC VIEW than previously noted by von Szalay (1998) and von Szalay and McConnaughey (2002). The QTC IMPACT method of stacking multiple pings could potentially ameliorate the influence of slope-affected pings, but it could also create new substrate classes by combining normal and non-normal reflections. Although it is presumed that QTC IMPACT software could distinguish among substrates at a constant depth in a flat seafloor area with no vessel pitch and roll, this type of situation is not realistic for the NMFS bottom trawl surveys in the Gulf of Alaska and Aleutians Islands. If steep seafloor areas can be distinguished from vessel motion through careful incorporation of vessel motion measurements, slope may be considered as a substrate modifier or as a significantly different substrate, depending on the species of interest for which habitat is being defined.

Assumption one: scale of measurements

The QTC IMPACT method of PCA (without standardization of data) results in a higher amount of variance explained because it is based on a few variables with the highest variance, which are also highly correlated. Those EMs without much variance only make a minor contribution to the PCA solution; however, it remains unclear whether forcing EMs to vary through standardization—a process that could possibly include both discriminating (signal) and nondiscriminating measures (noise) of echo energy, timespread, and skewness (van Walree et al., 2005)—increases or decreases statistical power for discriminating substrate types. The user is left having to choose between conducting a nonstandardized PCA where nearly all variables are collinear or conducting a standardized PCA that may be based mostly on noise. Including fully collinear (e.g., the sums of) variables and correlated variables in a PCA does not provide additional discriminatory information, but it does change the results. Therefore users may find it beneficial to conduct an additional PCA without these collinear variables and determine how much the substrate groupings change. Perhaps not coincidentally, our findings that only three to six variables within each of the acoustic data sets were somewhat independent (provided discriminatory power) matches well with van Walree et al.'s (2005) description of six acoustic algorithms and Kloser et al.'s (2001) description of four algorithms. Our results indicate that the EMs are not standardized before PCA, and because this is not mentioned in the literature, it may be an unexpected problem for users. The lack of standardization among collinear and correlated variables might partly explain why QTC IMPACT software typically requires only three eigenvectors to explain more than 95% of covariance (Ellingsen et al., 2002; Legendre et al., 2002).

Assumption two: first echo

Surveys conducted in shallow water with high sampling rates and surveys conducted in deep water with low sampling rates (such as that of the NMFS 2003 FV *Gladiator*) are equally vulnerable to accidentally including second, or later, seafloor reflections in QTC IMPACT analysis. For example, the RV *Rangithi* 1999 and RV *Pallasi* 2004 data sets both required 9.4 m below the start of the first seafloor reflection to achieve 251 samples, but the range of these data sets was shallower (Table 1). Because both of these data sets were recorded directly into QTC VIEW, the raw data could not be examined to determine if additional echoes were recorded or not. However, both data sets had characteristic drops in the values of the first group of EMs between approximately 9 and 4 m depth, indicating a probable increasing inclusion of the second echo with a decrease in depth (see Fig. 3). By 4 m in depth, both data sets should have included most of the second echo. The sharp increase in EM 1 just below 3 m in the RV *Pallasi* 2004 data set, at the shallowest depth, may indicate partial inclusion of the third echo. Thus accidental analysis of more than one echo with QTC IMPACT can cause strong depth-related influences and can create significantly different echogram measurements such that additional substrate classes could be created. To avoid such problems, users need to compare the depth range for their echogram measurement analysis (echo envelope) to the range of depths in their study area.

Conclusions

The need for a cost-effective approach to classify seafloor substrates, in order to define EFH across areas such as the NMFS bottom trawl surveys of the Gulf of Alaska and Aleutian Islands, remains strong. Because of the unexpected problems with the QTC IMPACT processing steps and creation of EMs, it seems highly likely that QTC IMPACT users are producing substrate classifications based on problems implementing the software or analyzing the measurements. Although data-gathering or data-processing errors are common across all such analyses, there is little chance to correct such errors when using a black box system. Therefore for future projects more transparent analytical methods will be needed, such as the published algorithms in Kloser et al. (2001) and van Walree et al. (2005), for translating acoustic data into EFH.

Acknowledgments

We thank D. Urban for supplying the RV *Resolution* 2003 data, I. Murfitt for the RV *Pallasi* 2004 data,

C. Grandin for the C.C.G.S. RV *John P. Tully* 2002 data, and J. Hewitt for the RV *Rangithi* 1999 data. The data contributors also provided helpful manuscript reviews, along with D. Somerton, R. McConnaughey, S. Syrjala, L. Bonacci, and three anonymous reviewers. S. Syrjala, P. Spencer, and others provided helpful statistical advice. The video and multibeam groundtruth data from the Aleutian Islands were collected with support from the North Pacific Research Board.

Literature cited

Anderson, J. T.
2001. Classification of marine habitats using submersible and acoustic seabed techniques. *In* Spatial processes and management of fish populations (G. H. Kruse, N. Bez, A. Booth, M. W. Dorn, S. Hills, R. N. Lipcius, D. Pelletier, C. Roy, S. J. Smith, and D. Witherells, eds.), p. 377–393. Alaska Sea Grant Rep., Alaska Sea Grant Program, Univ. Alaska Fairbanks. AK-SG-01-02.

Anderson, J. T., R. S. Gregory, and W. T. Collins.
2002. Acoustic classification of marine habitats in coastal Newfoundland. ICES J. Mar. Sci. 59:156–167.

Ellingsen, K. E., J. S. Gray, and E. Bjørnbom.
2002. Acoustic classification of seabed habitats using the QTC VIEW^{TM} system. ICES J. Mar. Sci. 59: 825–835.

Federal Register.
2002. Magnuson-Stevens Act Provisions; Essential Fish Habitat (EFH). 50 CFR Part 600, Federal Register 67(12):2343–2383. Office of the Federal Register, National Archives and Records Administration (NARA), College Park, MD.

Kloser, R. J., N. J. Bax, T. Ryan, A. Williams, and B. A. Barker.
2001. Remote sensing of seabed types in the Australian South East Fishery; development and application of normal acoustic techniques and associated "ground truthing." Mar. Freshw. Res. 52:475–489.

Legendre, P.
2003. Reply to the comment by Preston and Kirlin on "Acoustic seabed classification: improved statistical method". Can. J. Fish. Aquat. Sci. 60:1301–1305.

Legendre, P., K. E. Ellingsen, E. Bjørnbom, and P. Casgrain.
2002. Acoustic seabed classification: improved statistical method. Can. J. Fish. Aquat. Sci. 59:1085–1089.

Legendre, P., and L. Legendre.
1998. Numerical ecology, 2nd ed., 853 p. Elsevier Science B.V., Amsterdam. [In English.].

Manly, B. F. J.
1994. Multivariate statistical methods, 2nd ed., 215 p. Chapman and Hall, New York, NY.

Morrison, M. A., S. F. Thrush, and R. Budd.
2001. Detection of acoustic class boundaries in soft sediment systems using the seafloor acoustic discrimination system QTC VIEW. J. Sea Res. 46:233–243.

Neter, J., W. Wasserman, and M. H. Kutner.
1990. Applied linear statistical models: regression, analysis of variance, and experimental designs, 3rd ed., 1181 p. Richard D. Irwin, Homewood, IL.

Orlowski, A.
1984. Application of multiple echoes energy measurements for evaluation of sea bottom type. Oceanologia 19:61–78.

Pace, N. G., and R. V. Ceen.
1982. Seabed classification using the backscattering of normally incident broadband acoustic pulses. Hydrogr. J. 26:9–16.

Preston, J. M., and R. L. Kirlin.
2003. Comment on "Acoustic seabed classification: improved statistical method". Can. J. Fish. Aquat. Sci. 60:1299–1300.

Rooper, C. N., and M. Zimmermann.
2007. A bottom-up methodology for integrating underwater video and acoustic mapping for seafloor substrate classification. Cont. Shelf Res. 27:947–957.

van Walree, P. A., J. T. Tegowski, C. Laban,, and D. G. Simons.
2005. Acoustic seafloor discrimination with echo shape parameters: A comparison with the ground truth. Cont. Shelf Res. 25:2273–2293.

von Szalay, P. G.
1998. The feasibility of using single beam seabed classification systems to identify and quantify slope rockfish habitat in the Gulf of Alaska. M.S. thesis, 158 p. Univ. Washington, Seattle, WA.

von Szalay, P. G., and R. A. McConnaughey.
2002. The effect of slope and vessel speed on the performance of a single beam acoustic seabed classification system. Fish. Res. 56:99–112.